ARL

US Army Research Laboratory

Optical Flow-Based State Estimation for Guided Projectiles

by Moshe Hamaoui

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

US Army Research Laboratory

Optical Flow-Based State Estimation for Guided Projectiles

by Moshe Hamaoui
Weapons and Materials Research Directorate, ARL

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
June 2015	Final	1 September 2014–28 April 2015

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Optical Flow-Based State Estimation for Guided Projectiles	
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Moshe Hamaoui	
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
US Army Research Laboratory ATTN: RDRL-WML-F Aberdeen Proving Ground, MD 21005-5069	ARL-TR-7321

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Optical flow-based state estimation could significantly enhance the scope and performance of future guided projectiles by mitigating reliance on Global Positioning Systems (GPS) and augmenting traditional inertial measurement units and other navigational sensors. The goal of this work is to specialize and apply the theory of optical flow geometry to projectile vision-based navigation. This report describes the construction of an optical flow (OF) simulation tool for projectile flight and presents the equations and constraints relating OF to projectile motion. Specifically, the differential epipolar constraint is used to recover the body-fixed translational velocity vector (up to scale) and angular rates from the projected flow field. Finally, an example simulation is presented wherein we generate vehicle state and OF "ground truth" histories, corresponding synthetic video sequence, measured OF, and state estimation history.

15. SUBJECT TERMS

optical flow, vision-based navigation, epipolar constraint, guided munitions

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Moshe Hamaoui
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	28	19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified			410-306-0968

Contents

List of Figures

List of Tables

1. Introduction

State awareness is an essential prerequisite for autonomous navigation. Like their biological counterparts, modern guided munitions rely on inertial and environmental measurements to infer pose, position, and motion with respect to their surroundings. The Earth's magnetic field, position of the sun, global positioning system (GPS), and even the ground-sky temperature gradient are some of the environmental cues that have been leveraged toward projectile Guidance Navigation and Control (GNC). Only recently, however, have researchers begun to explore and develop vision-based techniques to aid projectile GNC. This report focuses specifically on the use of optical flow (OF) in recovering vehicle state.

Just as points and objects in a 3-dimensional (3-D) scene are projected onto a plane to form an image, so are velocity vectors of these same points and objects. This is the projected OF field. While the flow field can be described in closed form, the inverse problem of recovering self-motion from OF is an area of ongoing research. Several literature reviews describe the theory and techniques used to infer ego-motion from OF.[1–3] The mathematics relating self-motion to perceived OF, developed during the past 30 years, has been applied notionally to various experimental platforms, including robots,[4] unmanned aerial vehicles (UAVs),[5] aircraft,[6] and missiles.[7]

In general, the possibility of motion within a nonrigid scene adds an unknown number of degrees of freedom. Fortunately, the projectile application typically admits several reasonable assumptions and constraints that make the problem more tractable. In particular, if the visible scene is rigid, then the OF can be described entirely in terms of the motion of the observer and camera intrinsic parameters (e.g., focal length, distortion). Under this assumption, we can make use of the so-called continuous differential epipolar constraint[8] to solve for the body-fixed translational velocity (up to scale[9]) and angular rates.

The goal of this work is to specialize and apply the theory of OF geometry to projectile vision based navigation. In the following, we describe the construction of an OF simulation tool for projectile flight. We then present the equations and constraints relating OF to self-motion. Finally, an example simulation is presented wherein we generate vehicle state and OF "ground truth" histories, corresponding video sequence, measured OF, and state estimation history. The report concludes with a discussion of results and future research paths.

2. Conventions

1) Vectors are bold lowercase (e.g., $\mathbf{v}$). Matrices are bold uppercase italic (e.g., $\boldsymbol{R}$). Scalars are lowercase italic (e.g., f).

2) Vectors with homogeneous coordinates appear under the tilde (e.g., $\tilde{\mathbf{v}}$).

3) $[\mathbf{a}]_\times$ designates the skew-symmetric square matrix associated with the vector $\mathbf{a}$ so that $[\mathbf{a}]_\times\mathbf{b} = \mathbf{a} \times \mathbf{b}$

$$[\mathbf{a}]_\times \equiv \begin{bmatrix} 0 & -a_k & a_j \\ a_k & 0 & -a_i \\ -a_j & a_i & 0 \end{bmatrix}. \tag{1}$$

4) There are several coordinates systems used, defined by location of origin followed by orthonormal basis vectors.

   a. Earth frame, $\Gamma_E = \{\mathbf{e}_O, \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$. This is the gun-fixed inertial frame with origin $\mathbf{e}_O$ coincident with the firing platform. We define positive $\hat{\mathbf{z}}$ pointing toward the center of the Earth, $\hat{\mathbf{x}}$ points downrange, and $\hat{\mathbf{y}}$ completes the right-handed basis.

   b. Body frame, $\Gamma_B = \{\mathbf{b}_O, \hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$. This projectile-fixed inertial frame is instantaneously at rest with respect to Γ_E. The origin $\mathbf{b}_O$ is coincident with the optical center of the projectile-mounted camera, a distance f behind the focal plane (Fig. 1). $\hat{\mathbf{i}}$ points along the projectile's long axis and defines the optical axis. $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ span the focal plane.
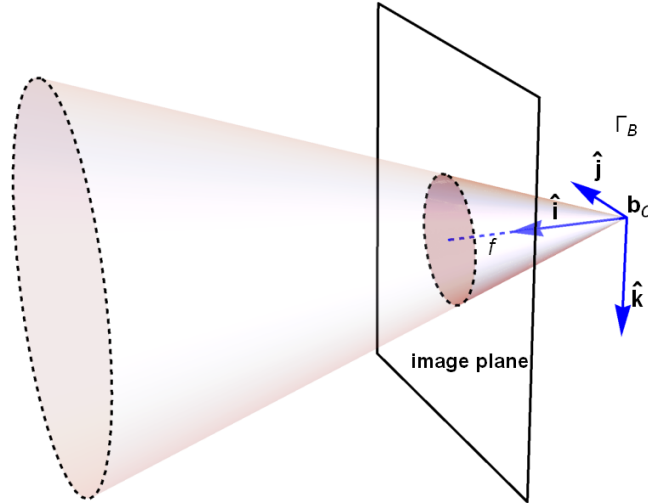

Fig. 1   Body frame definition in relation to image plane

    c. Rotating Body frame, $\boldsymbol{\Gamma}_{B'}$. This noninertial frame is coincident with Γ_B, but is subject to both translational and rotational differential motion (with respect to Γ_E). Practically, this is the frame from which OF is measured.

5) When viewed from a particular frame, a vector or vector component may be followed by a superscript denoting the reference frame (e.g., $\mathbf{v}^B$ or v_j^B). If clear from the context, the superscript may be omitted.

6) A matrix transformation between frames will be designated by subscript and superscript (e.g., R_E^B designates a rotation from Γ_E to Γ_B), unless clear from the context.

3. Projectile Vision Model

State estimation performance will depend on imaging constraints, flight behavior, scene structure, and OF measurement methodology. Table 1 lists some of these parameters.

Table 1   Projectile vision model parameters

Input			Output
Projectile	**Camera**	**Environment**	
• Quadrant Elevation (QE) • Initial velocity • roll rate • max pitch/yaw perturbations • peak angular rates	• frame rate • focal length • field of view • imaging resolution	• Scene structure (point cloud)	• 12 state history • Imaging history • Optical-flow history

In order to develop an OF simulation tool, we must formulate the component models for projectile flight, imaging, and OF in terms of these parameters. The remainder of this section details the construction and integration of these models.

3.1 Flight Model

For vision-based applications, we are concerned with the image-space response to specified flight dynamics rather than with aerodynamic forces and moments. We therefore adopt a simple flight model, which extends the point mass model by imposing prescribed pose-perturbations about the velocity vector.

The translational motion is given by the point mass vacuum model:

$$\mathbf{v}^E[t] = \begin{bmatrix} v_0\cos[QE]\cos[\psi] \\ v_0\cos[QE]\sin[\psi] \\ gt - v_0\sin[QE] \end{bmatrix}, \tag{2}$$

$$\mathbf{r}^E[t] = \begin{bmatrix} t\cos[QE]\cos[\psi]v_0 + x_0^E \\ t\cos[QE]\sin[\psi]v_0 + y_0^E \\ \dfrac{gt^2}{2} - t\sin[QE]v_0 + z_0^E \end{bmatrix}. \tag{3}$$

The projectile's long axis ($\hat{\mathbf{i}}$) is nominally oriented along $\mathbf{v}$. The pitch and yaw angles associated with $\mathbf{v}$ are, respectively:

$$\theta_{\mathbf{v}} = \sin^{-1}\left[-\frac{\mathbf{v}\cdot\hat{\mathbf{z}}}{\|\mathbf{v}\|}\right], \tag{4}$$

$$\psi_{\mathbf{v}} = \tan^{-1}\left[\frac{\mathbf{v}\cdot\hat{\mathbf{y}}}{\mathbf{v}\cdot\hat{\mathbf{x}}}\right]. \tag{5}$$

The pitch perturbation θ_p is taken to be a sinusoidal function subject to:

$$\dot{\theta}_p = \left(\dot{\theta}_p\right)_{max}\cos[\gamma t], \tag{6}$$

$$\theta_p[0] = 0. \tag{7}$$

Where $\left(\dot{\theta}_p\right)_{max}$ and γ are constants. The solution to Eq. 6 is:

$$\theta_p = \frac{\left(\dot{\theta}_p\right)_{max}}{\gamma}\sin[\gamma t]. \tag{8}$$

From Eq. 8, it can be seen that the prefactor $\frac{\left(\dot{\theta}_p\right)_{max}}{\gamma}$ is the maximum pitch deflection angle $\left(\theta_p\right)_{max}$ from $\mathbf{v}$ so that

$$\theta_p = \left(\theta_p\right)_{max}\sin\left[\frac{\left(\dot{\theta}_p\right)_{max}}{\left(\theta_p\right)_{max}}t\right], \tag{9}$$

$$\dot{\theta}_p = \left(\dot{\theta}_p\right)_{max}\cos\left[\frac{\left(\dot{\theta}_p\right)_{max}}{\left(\theta_p\right)_{max}}t\right]. \tag{10}$$

Similarly, for yaw perturbation, we have

$$\psi_p = \left(\psi_p\right)_{max}\sin\left[\frac{\left(\dot{\psi}_p\right)_{max}}{\left(\psi_p\right)_{max}}t\right], \tag{11}$$

$$\dot{\psi}_p = \left(\dot{\psi}_p\right)_{max} \cos\left[\frac{\left(\ddot{\psi}_p\right)_{max}}{\left(\dot{\psi}_p\right)_{max}} t\right]. \tag{12}$$

The roll angle ϕ will depend on a prescribed roll rate $\dot{\phi}$ according to

$$\phi = \dot{\phi}t. \tag{13}$$

From Eqs. 4, 5 and 9–13 we construct the Euler angles:

$$\begin{bmatrix} \phi \\ \theta \\ \psi \end{bmatrix} = \begin{bmatrix} 0 \\ \sin^{-1}\left[-\frac{\mathbf{v}\cdot\hat{\mathbf{z}}}{\|\mathbf{v}\|}\right] \\ \tan^{-1}\left[\frac{\mathbf{v}\cdot\hat{\mathbf{y}}}{\mathbf{v}\cdot\hat{\mathbf{x}}}\right] \end{bmatrix} + \begin{bmatrix} \dot{\phi}t \\ \left(\dot{\theta}_p\right)_{max} \sin\left[\frac{\left(\ddot{\theta}_p\right)_{max}}{\left(\dot{\theta}_p\right)_{max}} t\right] \\ \left(\dot{\psi}_p\right)_{max} \sin\left[\frac{\left(\ddot{\psi}_p\right)_{max}}{\left(\dot{\psi}_p\right)_{max}} t\right] \end{bmatrix}. \tag{14}$$

Assuming $\psi_{\mathbf{v}}$ and $\dot{\phi}$ are constant in time and $v \equiv \|\mathbf{v}\|$, the angular rates are:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 0 \\ \dfrac{(v\,\dot{\mathbf{v}} - \dot{v}\,\mathbf{v})\cdot\hat{\mathbf{z}}}{v^2\sqrt{1 - \left(\frac{\mathbf{v}\cdot\hat{\mathbf{z}}}{v}\right)^2}} \\ 0 \end{bmatrix} + \begin{bmatrix} \dot{\phi} \\ \left(\dot{\theta}_p\right)_{max} \cos\left[\frac{\left(\ddot{\theta}_p\right)_{max}}{\left(\dot{\theta}_p\right)_{max}} t\right] \\ \left(\dot{\psi}_p\right)_{max} \cos\left[\frac{\left(\ddot{\psi}_p\right)_{max}}{\left(\dot{\psi}_p\right)_{max}} t\right] \end{bmatrix}. \tag{15}$$

Using Eqs. 14 and 15 we can generate example 12-state histories with prescribed angular characteristics. $\{\dot{\theta}, \dot{\psi}\}$ correspond to the effective slew rates of the strap-down imager. In Eq. 15 the first term in $\dot{\theta}$ accounts for the rate of ballistic overturn. Thus, while $\dot{\psi}$ is directly specified by $\dot{\psi}_p$, $\dot{\theta}$ will differ slightly from the specified $\dot{\theta}_p$.

3.2 State Transformations

Euler angles define the relation between the Earth and body frames. It will be useful to transform the generated states from Earth to body frame, and vice versa. The rotation sequence used here is the so-called ZYX Aerospace sequence[10,11]

$$\begin{aligned} R_E^B &= R_\phi^x R_\theta^y R_\psi^z \\ &= \begin{bmatrix} c_\theta c_\psi & c_\theta s_\psi & -s_\theta \\ c_\psi s_\theta s_\phi - c_\phi s_\psi & c_\phi c_\psi + s_\theta s_\phi s_\psi & c_\theta s_\phi \\ c_\phi c_\psi s_\theta + s_\phi s_\psi & -c_\psi s_\phi + c_\phi s_\theta s_\psi & c_\theta c_\phi \end{bmatrix}. \end{aligned} \tag{16}$$

The rows of R_E^B define the orthonormal basis of the body frame with respect to $\Gamma_E, \{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}^E$. The inverse rotation is $R_B^E = (R_E^B)^{\mathrm{T}}$, and the body-fixed translational velocity $\mathbf{v}^B = \{v_i, v_j, v_k\}$ is then

$$\mathbf{v}^B = R_E^B \mathbf{v}^E . \tag{17}$$

The body-fixed angular rates $\boldsymbol{\omega}^B = \{\omega_i, \omega_j, \omega_k\}$ are given by:[10]

$$\boldsymbol{\omega}^B = \begin{bmatrix} 1 & 0 & -s_\theta \\ 0 & c_\phi & c_\theta s_\phi \\ 0 & -s_\phi & c_\theta c_\phi \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} . \tag{18}$$

The pitch angle of attack, α, and the yaw angle of attack, β, are defined as

$$\alpha = \tan^{-1}\left[\frac{v_k}{v_i}\right] , \tag{19}$$

$$\beta = \sin^{-1}\left[\frac{v_j}{\|\mathbf{v}\|}\right] . \tag{20}$$

3.3 OF Field

We next seek to define the flow field as a function of vehicle state. The imager is taken to be a strap-down forward-facing pinhole camera. To be specific, we imagine a scene point $\mathbf{p} \in \mathbb{R}^3$ and corresponding projection $\mathbf{q} \in \mathbb{R}^3$ on the focal plane.[12] Our goal is to find the transformation from $\mathbf{p}^E$ to $\mathbf{q}^B$ and construct the OF field $\{\mathbf{q}^B, \dot{\mathbf{q}}^B\}$.

3.3.1 Projected Scene

To project a scene point defined in Γ_E onto the focal plane, several sequential transformations are needed: 1) translation to body origin, 2) rotation to body frame, and 3) projection to focal plane. For efficient matrix representation, we proceed using homogenous coordinates and corresponding matrix transformations. The homogenous coordinates of a point $\{x, y, z\}$ in $\mathbb{R}^3$ maps as

$$\{x, y, z\} \rightarrow \{x, y, z, 1\} . \tag{21}$$

Translation. The Γ_E frame is first translated by the displacement vector $\mathbf{r}[t]$ from $\mathbf{e}_O$ to $\mathbf{b}_O$:

$$T = \begin{bmatrix} 1 & 0 & 0 & -r_x \\ 0 & 1 & 0 & -r_y \\ 0 & 0 & 1 & -r_z \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{22}$$

Rotation. We next rotate the reference basis into the body basis using

$$R = \begin{bmatrix} & & & 0 \\ & R_E^B & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{23}$$

Projection. For a point viewed in the body frame, the pinhole camera model defines the projection mapping from $\mathbf{p}^B$ to $\mathbf{q}^B$, as illustrated in Fig. 2.
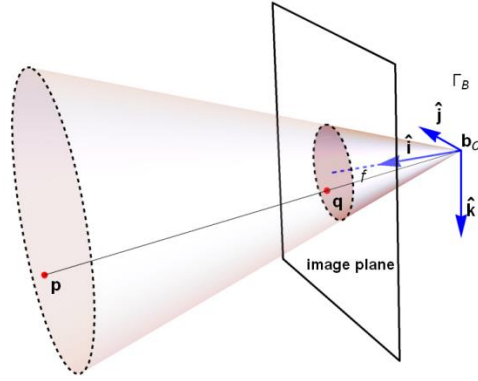

Fig. 2   Pinhole camera projection

Retaining the depth coordinate in $\mathbf{q}$, we then have:

$$q_i^B = f, \qquad q_j^B = f\frac{p_j^B}{p_i^B}, \qquad q_k^B = f\frac{p_k^B}{p_i^B}. \tag{24}$$

In matrix form, the projective transformation is then

$$P = \frac{f}{p_i^B}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/f & 0 & 0 & 0 \end{bmatrix}. \tag{25}$$

Where p_i^B is the $\hat{\mathbf{i}}$ component of $\widetilde{\mathbf{p}}^B = RT\widetilde{\mathbf{p}}^E$. The complete transformation for a point $\widetilde{\mathbf{p}}^E$ is:

$$\widetilde{\mathbf{q}}^B = PRT\widetilde{\mathbf{p}}^E. \tag{26}$$

Where $\widetilde{\mathbf{q}}$ is of the form $\widetilde{\mathbf{q}} = \{f, q_j, q_k, 1\}$, and only $\{q_j, q_k\}$ are needed to form the image in the focal plane.

3.3.2 Projected Velocity Field

As we shall now introduce motion effects into the projection model, we will need to distinguish between rotating and nonrotating body frames. The primed body

frame $\Gamma_{B'}$ shall be the rotating (and translating) frame, while the unprimed Γ_B is instantaneously at rest with respect to Γ_E and coincident with $\Gamma_{B'}$. The projected velocity is measured in the rotating frame and is found by differentiating the pinhole camera Eq. 24, leading to:[13]

$$
\begin{aligned}
\dot{q}_j^{B'} &= \left(-\frac{f_j p_j \dot{p}_i}{p_i^2} + \frac{f_j \dot{p}_j}{p_i} \right)^{B'} \\
\dot{q}_k^{B'} &= \left(-\frac{f_k p_k \dot{p}_i}{p_i^2} + \frac{f_k \dot{p}_k}{p_i} \right)^{B'} .
\end{aligned}
\tag{27}
$$

While the position vector is simply $\mathbf{p}^{B'} = \mathbf{p}^B$, the derivative $\dot{\mathbf{p}}^{B'}$ will depend on the motion of the viewing platform. The derivative of a vector measured in an inertial frame is related to the corresponding derivative in the co-located rotating frame by:[14,15]

$$
\dot{\mathbf{p}}^B = \dot{\mathbf{p}}^{B'} + \boldsymbol{\omega} \times \mathbf{p} + \mathbf{v}^B .
\tag{28}
$$

Where $\mathbf{v}^B = \dot{\mathbf{b}}_O'^{\,B}$ is the translational velocity of the projectile with respect to Γ_B, and $\boldsymbol{\omega}$ is the angular velocity of $\Gamma_{B'}$ with respect to Γ_B. Noting that $\dot{\mathbf{p}}^B = 0$ and rearranging Eq. 28 gives the time derivative of a scene point $\mathbf{p}$ as viewed from the projectile:

$$
\dot{\mathbf{p}}^{B'} = -\mathbf{v}^B - \boldsymbol{\omega} \times \mathbf{p} ,
\tag{29}
$$

$$
\begin{aligned}
\dot{p}_i^{B'} &= \left(-v_i - p_k \omega_j + p_j \omega_k \right)^B \\
\dot{p}_j^{B'} &= \left(-v_j + p_k \omega_i - p_i \omega_k \right)^B \\
\dot{p}_k^{B'} &= \left(-v_k - p_j \omega_i + p_i \omega_j \right)^B
\end{aligned}
\tag{30}
$$

Inserting Eq. 30 into Eq. 27, we then have:[4]

$$
\begin{aligned}
\dot{q}_j &= \frac{f_j p_j v_i}{p_i^2} - \frac{f_j v_j}{p_i} + \frac{f_j p_k \omega_i}{p_i} + \frac{f_j p_j p_k \omega_j}{p_i^2} + \left(-f_j - \frac{f_j p_j^2}{p_i^2} \right) \omega_k \\
\dot{q}_k &= \frac{f_k p_k v_i}{p_i^2} - \frac{f_k v_k}{p_i} - \frac{f_k p_j \omega_i}{p_i} + \left(f_k + \frac{f_k p_k^2}{p_i^2} \right) \omega_j - \frac{f_k p_j p_k \omega_k}{p_i^2} .
\end{aligned}
\tag{31}
$$

Which can be expressed in matrix form as

$$
\begin{bmatrix} \dot{q}_j \\ \dot{q}_k \end{bmatrix} = \begin{bmatrix} \dfrac{f_j p_j}{p_i^2} & -\dfrac{f_j}{p_i} & 0 & \dfrac{f_j p_k}{p_i} & \dfrac{f_j p_j p_k}{p_i^2} & -f_j - \dfrac{f_j p_j^2}{p_i^2} \\[2ex] \dfrac{f_k p_k}{p_i^2} & 0 & -\dfrac{f_k}{p_i} & -\dfrac{f_k p_j}{p_i} & f_k + \dfrac{f_k p_k^2}{p_i^2} & -\dfrac{f_k p_j p_k}{p_i^2} \end{bmatrix} \begin{bmatrix} v_i \\ v_j \\ v_k \\ \omega_i \\ \omega_j \\ \omega_k \end{bmatrix}. \tag{32}
$$

Together, Eqs. 26 and 32 give the OF field $\{\mathbf{q}, \dot{\mathbf{q}}\}$ as a function scene structure and projectile motion. For simulation purposes, $\mathbf{v}^B$ and $\boldsymbol{\omega}^B$ are calculated respectively by Eqs. 17 and 18, while $\widetilde{\mathbf{p}}^B = RT\widetilde{\mathbf{p}}^E$.

4. State Estimation

4.1 Direct Differentiation of Perspective Projection

Eq. 32 shows the relation between projected velocity field and the 6 body rates composing the linear and angular velocity vectors. The form of this equation, which directly references the point location $\mathbf{p}$, is useful for simulation of the flow vectors for a given (known) scene structure and vehicle state history. For purposes of state estimation where scene structure is not typically known a-priori, we can exchange $\{p_j, p_k\}$ in favor of the (measurable) image coordinates $\{q_j, q_k\}$. Rearranging the pinhole camera Eqs. 24 we have

$$
p_j = \frac{p_i q_j}{f_j}, \quad p_k = \frac{p_i q_k}{f_k}. \tag{33}
$$

Substituting Eq. 33 into Eq. 31, we have

$$
\dot{q}_j = \frac{q_j v_i}{p_i} - \frac{f_j v_j}{p_i} + \frac{f_j q_k \omega_i}{f_k} + \frac{q_j q_k \omega_j}{f_k} + \left(-f_j - \frac{q_j^2}{f_j}\right)\omega_k
$$
$$
\dot{q}_k = \frac{q_k v_i}{p_i} - \frac{f_k v_k}{p_i} - \frac{f_k q_j \omega_i}{f_j} + \left(f_k + \frac{q_k^2}{f_k}\right)\omega_j - \frac{q_j q_k \omega_k}{f_j}. \tag{34}
$$

The matrix analog of Eq. 32 is then

$$
\begin{bmatrix} \dot{q}_j \\ \dot{q}_k \end{bmatrix} = \begin{bmatrix} \dfrac{q_j}{p_i} & -\dfrac{f_j}{p_i} & 0 & \dfrac{f_j q_k}{f_k} & \dfrac{q_j q_k}{f_k} & -f_j - \dfrac{q_j^2}{f_j} \\[2ex] \dfrac{q_k}{p_i} & 0 & -\dfrac{f_k}{p_i} & -\dfrac{f_k q_j}{f_j} & f_k + \dfrac{q_k^2}{f_k} & -\dfrac{q_j q_k}{f_j} \end{bmatrix} \begin{bmatrix} v_i \\ v_j \\ v_k \\ \omega_i \\ \omega_j \\ \omega_k \end{bmatrix}. \tag{35}
$$

Each nth OF observation $\{\mathbf{q}, \dot{\mathbf{q}}\}_n$ then corresponds to 2 equations in 7 unknowns (6 body states plus the scene-point depth $p_{i,n}$). For N points, we have $2N$ equations and $6 + N$ unknowns. A minimal solution would then require 6 points (12 equations in 12 unknowns). It turns out, however, that $\mathbf{v}^B$ can only be solved for up to scale; we may solve for the direction, but not the magnitude of $\mathbf{v}^B$. This reflects the fact that scene depth is not directly observable from a single (monocular) observation; there is no way to decide whether the scene point is nearby and moving slowly or far away and moving quickly. In light of this state of affairs, we may as well set $p_{i,1} = 1$, which fixes the scale for the remaining (unknown) $p_{i,n}$. For N points, we now have $2N$ equations and $6 + (N - 1)$ unknowns. The exactly determined system requires 5 point observations.

Equation 35 is nonlinear, due to the appearance of point depth p_i in the denominator. The solution is not unique, and several real-valued solutions may exist. Further constraints are needed to identify the correct solution. Several reasonable constraints include:

1) $p_i > 0$. This assumes the field of view (FOV) is less than $180°$ and asserts that $\mathbf{p}$ cannot be seen behind the camera. This was implicit when fixing $p_{i,1} = 1$ but is true of all point depths $p_{i,n}$.

2) $v_i > 0$. This asserts that the projectile never stops (during flight) or travels tail-first (with respect to Γ_E).

3) $v_i^2 > v_j^2 \wedge v_i^2 > v_k^2$. This constrains the total angle of attack to be less than $45°$.[16]

Even with these constraints, however, our numerical trials often returned multiple real solutions, failing to provide a robust approach to state estimation. In the next section, we seek a stricter physical constraint to further narrow the solution space.

4.2 Differential Epipolar Constraint

Not all solutions to Eq. 35 are physically realizable. Geometric considerations lead to the so-called differential epipolar constraint,[1,8] relating the flow field $\{\mathbf{q}, \dot{\mathbf{q}}\}$ to the observer's motion $\{\mathbf{v}, \boldsymbol{\omega}\}$:

$$(\mathbf{q}^{\text{T}}[\mathbf{v}]_\times \dot{\mathbf{q}} + \mathbf{q}^{\text{T}}[\boldsymbol{\omega}]_\times [\mathbf{v}]_\times \mathbf{q})^B = 0. \tag{36}$$

Where $\mathbf{q}$ and $\dot{\mathbf{q}}$ are 3-vectors of the form $\mathbf{q} = [f, q_j, q_k]$, and $\dot{\mathbf{q}} = [0, \dot{q}_j, \dot{q}_k]$. Because Eq. 36 is independent of $\mathbf{p}$, the number of unknowns remains exactly 6, irrespective of the number of flow vectors considered. To solve for $\{\mathbf{v}, \boldsymbol{\omega}\}$, we therefore require a minimum of 6 point-observations. The velocity scale-ambiguity

is again manifest, so we seek only to recover the unit vector $\hat{\mathbf{v}} = \mathbf{v}/v$. Additional projectile constraints are again given by:

$$\begin{aligned}
v_i{}^2 + v_j{}^2 + v_k{}^2 &= 1 \\
v_i{}^2 &> v_j{}^2 \\
v_i{}^2 &> v_k{}^2 \\
v_i &> 0 \, .
\end{aligned} \tag{37}$$

5. Simulation

State estimation will rely on accurate OF measurement, as well as robust method of solution. Both steps represent potential error sources. It will therefore be useful to compare performance using true versus measured OF. The model developed in Section 3.3 will generate the true OF field for a given 12-state history. The measured flow is derived from a video sequence corresponding to the identical state history.[17] An example of the corresponding point cloud and solid geometry projections is shown in Fig. 3.
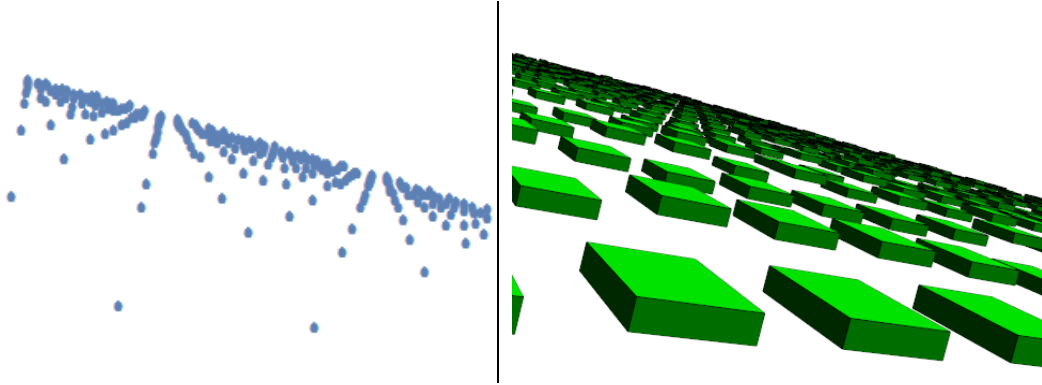
Fig. 3   **Projections of point cloud (left) and solid geometry onto the focal plane (right)**

The measurement of OF is a classical problem in computer vision and is a topic of continuing research. We proceed with a simple measurement scheme intended to reflect baseline rather than optimal performance. Let Δt be the time lapse between images. The SURF algorithm[18] is used to identify feature points $\{\mathbf{q}\}_M$ in successive image frames at time t and $t + \Delta t$. Then we may approximate the flow field as:

$$\begin{aligned}
\{\dot{\mathbf{q}}(t + \Delta t/2)\}_M &\approx \frac{\{\mathbf{q}(t + \Delta t)\}_M - \{\mathbf{q}(t)\}_M}{\Delta t} \, , \\
\{\mathbf{q}(t + \Delta t/2)\}_M &\approx \frac{\{\mathbf{q}(t + \Delta t)\}_M + \{\mathbf{q}(t)\}_M}{2} \, .
\end{aligned} \tag{38}$$

Equation 38 gives the measured flow field, analogous to the true flow field expressed in Eqs. 26 and 32. While every visible scene point has an associated true flow vector, only those feature points that can be tracked from frame to frame will have a measureable flow vector.

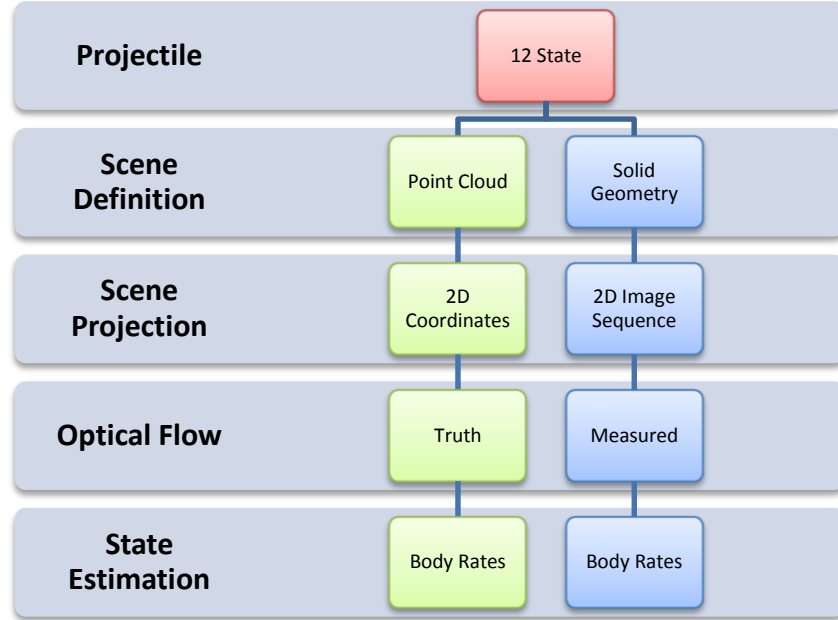Figure 4 describes the component models used in this simulation.

Fig. 4   **Simulation of projectile state estimation with exact or measured OF**

Table 2 lists the physical parameter definitions used in this example simulation.

Table 2   **Physical parameters used in simulation**

Quantity	Value	Units	Description
QE	15	deg	Elevation above horizon
v_0	100	m/s	Initial projectile speed
$\left(\theta_p\right)_{max}$	10	deg	Maximum pitch perturbation
$\left(\psi_p\right)_{max}$	15	deg	Maximum yaw perturbation
$\dot{\phi}$	2	Hz	Roll rate
$\left(\dot{\theta}_p\right)_{max}$	20	deg/s	Maximum pitch perturbation rate
$\left(\dot{\psi}_p\right)_{max}$	20	deg/s	Maximum yaw perturbation rate
FOV	35	deg	Field of view
Δt	0.02	s	Time lapse between images
res	300^2	pixels	Image resolution

12

6.   Results and Discussion

Arcing trajectories can limit the time during which useful imagery can be captured. While scene points are not constrained to reside in the ground plane, this is typically the neighborhood where rigid scenery will be more plentiful and rich. Thus, higher launch angle (QE) and smaller FOV will generally result in greater periods of scene occlusion. Figure 5 shows a plot of points-in-view as a function of time for this representative simulation.
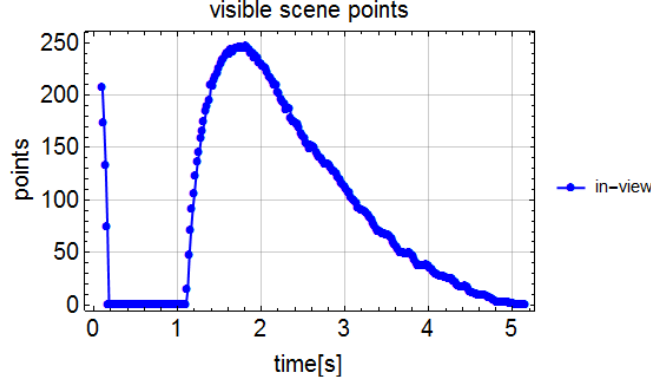

Fig. 5   Visible scene points during flight

At each time-step, the epipolar constraint is used to construct the least-squares minimization problem

$$\text{argmin} \sum_{n=1}^{N} [(\mathbf{q}_n^{\mathrm{T}}[\mathbf{v}]_\times \dot{\mathbf{q}}_n + \mathbf{q}_n^{\mathrm{T}}[\boldsymbol{\omega}]_\times [\mathbf{v}]_\times \mathbf{q}_n)^B]^2 \ . \tag{39}$$

Where N denotes the number of flow vectors considered, varying between 6 and 15 depending on availability. Equation 39 is further subject to the motional constraints of Eq. 37.[19]

The velocity vector $\mathbf{v}^B$, which can only be recovered up to scale, is used to calculate the corresponding angles of attack via Eqs. 19 and 20. Results for pitch and yaw angles of attack, as well as $\boldsymbol{\omega}$, are shown in Figs. 6 and 7.
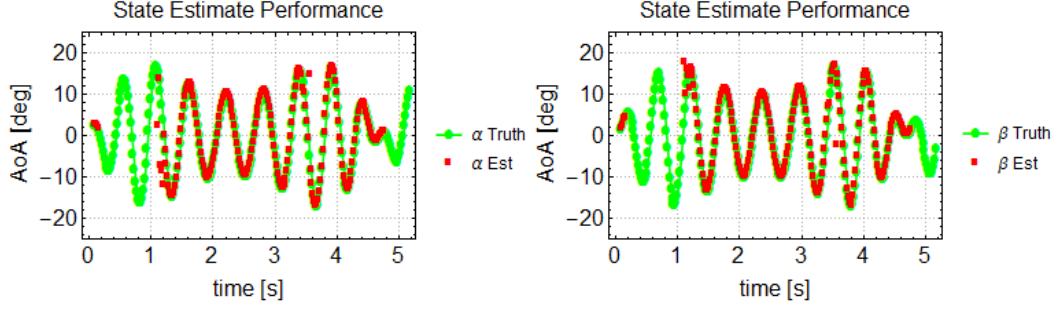
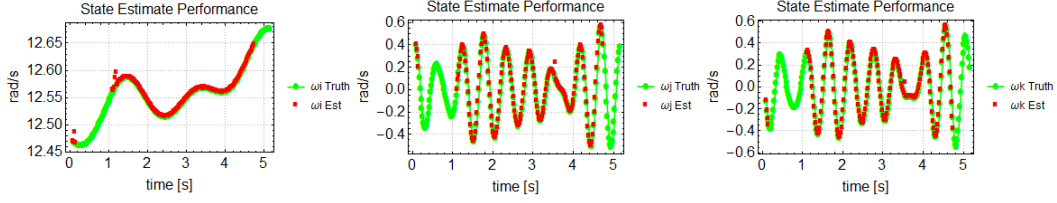**Fig. 6   Estimate of pitch and yaw angles of attack using perfect OF**

Fig. 7   Estimate of body-frame angular rates using perfect OF

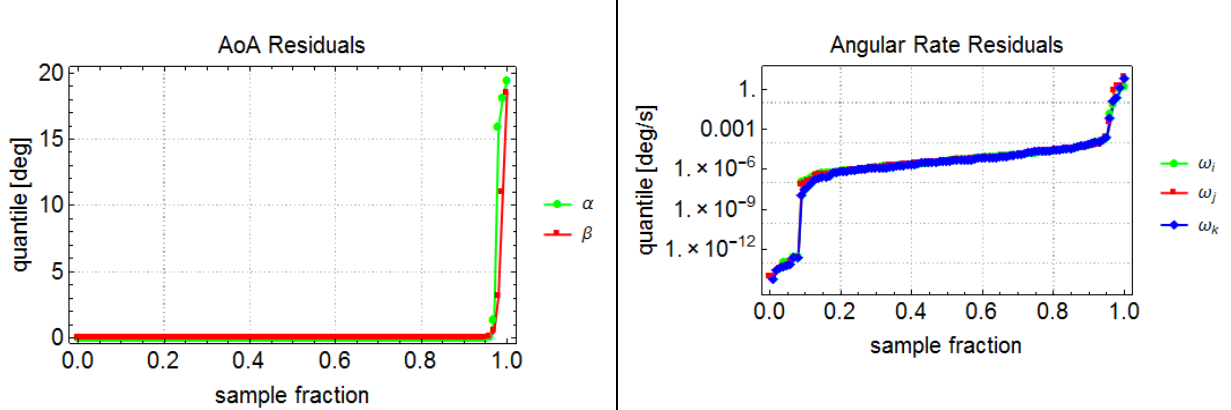The corresponding residuals are plotted by quantile in Fig. 8.

Fig. 8   Residuals of angles of attack and angular rates using perfect OF

It is interesting to compare these residuals to those for the identical case but without the stable-flight constraints of Eq. 37, as shown in Fig. 9.
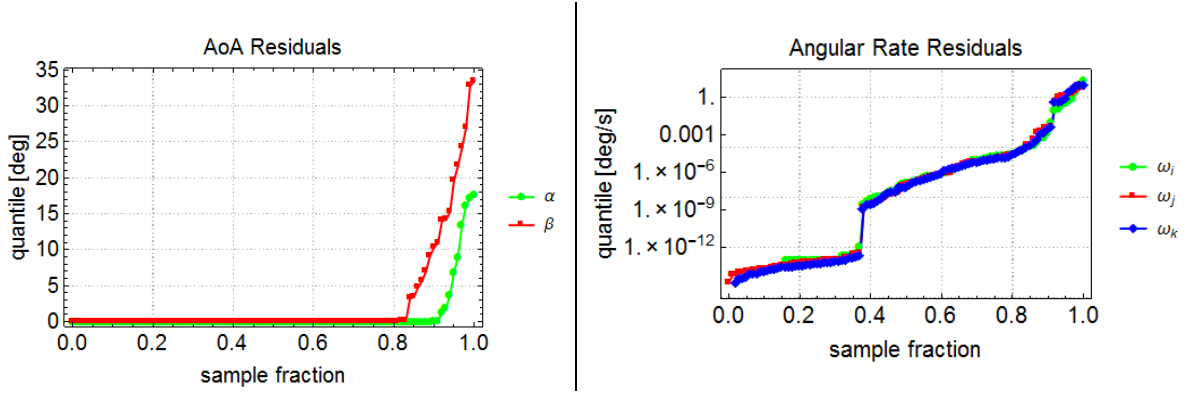
Fig. 9 **Estimate of body-frame angular rates using perfect OF without stable flight constraints**

For this example trajectory, these results suggest that for perfectly known OF, the epipolar constraint 36 in combination with projectile motion constraints 37 appear sufficient to precisely recover $\hat{\mathbf{v}}$ and $\boldsymbol{\omega}$.

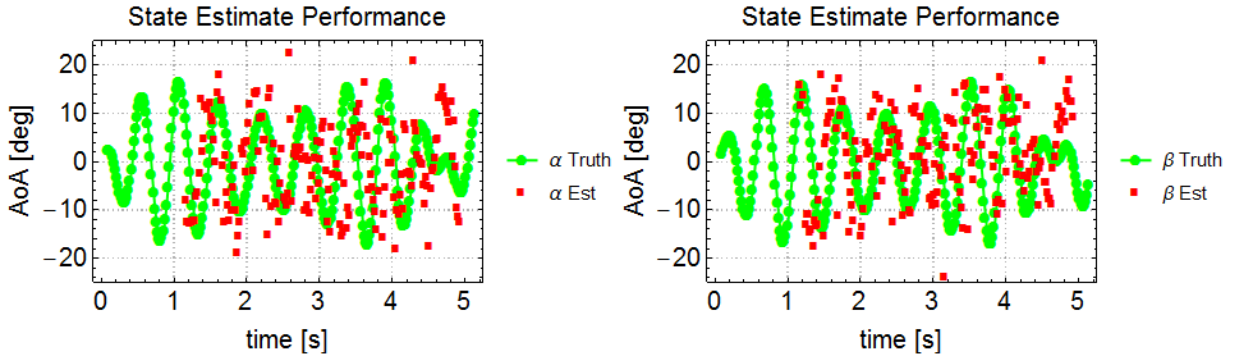Simulation results using measured OF are shown in Figs. 10 and 11. Residuals are shown in Fig. 12.

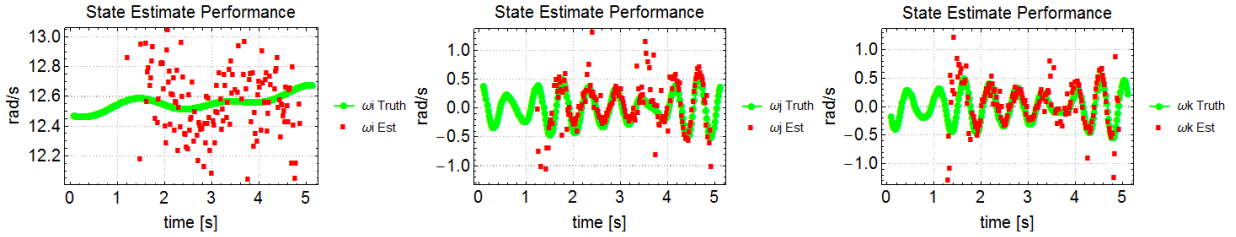**Fig. 10** **Estimate of pitch and yaw angles of attack using measured OF**

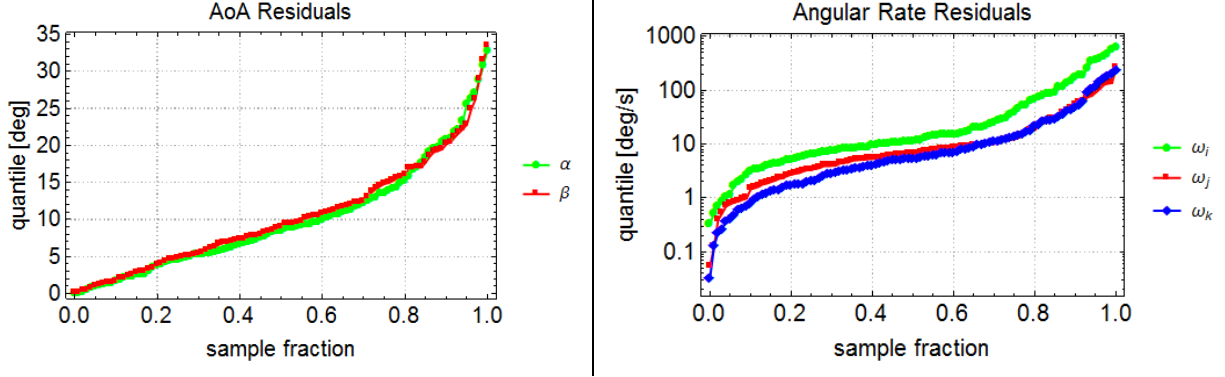**Fig. 11** **Estimate of body-frame angular rates using measured OF**

15

Fig. 12 Residuals of angles of attack and angular rates using measured OF

The increased estimation error when using measured OF is due to both limited imaging resolution, as well as mismatched frame-to-frame point correspondences (tracking error). The latter is easily identified, as in Fig. 13.
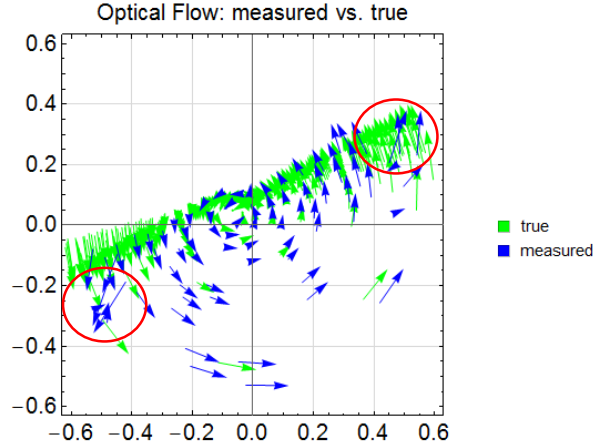

Fig. 13 Overlay of true vs. measured flow fields

These false flow vectors clearly violate the rigid-scene assumption. An automated outlier rejection scheme could be devised by demanding smooth flow variation across the image plane.

7. Conclusions and Future Work

Optical flow-based state estimation could significantly enhance the scope and performance of future guided projectiles by mitigating reliance on GPS and augmenting traditional inertial measurement units (IMUs) and other navigational sensors. Though not an ideal viewing platform, the projectile is nonetheless uniquely well-suited as an observer of OF. The characteristically elevated vantage

16

point and expansive view justify the rigid-scene assumption. Stable flight dynamics and forward motion also serve as state constraints.

The goal of this work has been to adapt the mathematics of OF geometry to projectile vision-based navigation and present the smart munitions community with an integrated modeling and simulation framework for OF-based state estimation. A flight-vision model has been presented that can be used to generate example 12-state histories with specified viewing behavior, along with corresponding synthetic imagery, exact OF, and measured OF.

The differential epipolar constraint was then used to recover translational and rotational velocity vectors $\hat{\mathbf{v}}$ and $\boldsymbol{\omega}$ from the projected flow field, and performance was evaluated against ground truth. State estimation was further improved by imposing stable-flight constraints. The approach was validated using exact OF and then applied to more realistic measured flow fields derived from a corresponding video sequence.

The precision of OF measurement, though not a primary focus of this study, will likely remain the primary error source. More sophisticated measurement techniques and/or dedicated OF-sensors will certainly result in improved estimation.

Another challenge will be achieving real-time performance. Speed of processing is limited by imaging resolution, frame-rate, and computing architecture, as well as measurement and solution algorithms. Fundamentally, the required (minimum) resolution and frame-rate will depend on the projectile's rate of motion (primarily $\dot{\phi}$) and range to scenery. A more quantified description of this trade-space will help to parametrically link imaging requirements to mission profile.

Finally, scene occlusion is particularly problematic for high-arcing trajectories and/or narrow FOV. In this study, the imager was assumed to be a bore-sighted monocular camera. If instead the camera were directed at some off-axis angle, projectile spin would allow for periodic state-updates along the entire trajectory. Likewise an array of 2 or more imagers could allow for more frequent updates and greater scene coverage.

8. References and Notes

1. Armangue X, Araujo H, Salvi J. A review on egomotion by means of differential epipolar geometry applied to the movement of a mobile robot. Pattern Recognition. 2003:2927–2944.

2. Raudies F, Neumann H. A review and evaluation of methods estimating ego-motion. Computer Vision and Image Understanding. 2012;116(5):606–633.

3. Corke P, Lobo J, Dias J. An introduction to inertial and visual sensing. The International Journal of Robotics Research. 2007;26(6):519–535.

4. Hutchinson S, Hager GD, Corke PI. A tutorial on visual servo control. Robotics and Automation, IEEE Transactions on. 1996;12(5):651–670.

5. Herisse B, Russotto F-X, Hamel T, Mahony R. Hovering flight and vertical landing control of a VTOL unmanned aerial vehicle using optical flow. In: International Conference on Intelligent Robots and Systems; IROS 2008. IEEE/RSJ, 2008.

6. Dusha D, Boles W. Walker R. Attitude estimation for a fixed-wing aircraft using horizon detection and optical flow. In: Digital Image Computing Techniques and Applications, 9th Biennial Conference of the Australian Pattern Recognition Society, 2007.

7. Manchester IR, Savkin AV, Faruqi FA. Method for optical-flow-based precision missile guidance. IEEE Transactions on Aerospace and Electronic Systems. 2008;44(3):835–851.

8. Brooks MJ, Chojnacki W, Baumela L. Determining the egomotion of an uncalibrated camera from instantaneous optical flow. JOSA A. 1997;14(10):2670–2677.

9. This provides the angles of attack in the body-fixed pitch and yaw planes.

10. Kuipers JB. Quaternions and rotation sequences: A primer with applications to orbits, aerospace and virtual reality. Princeton (NJ): Princeton University Press, 2002.

11. The notation R_ϕ^x denotes a rotation by an angle ϕ about the current x axis.

12. This could be described more concisely as an $\mathbb{R}^3 \to \mathbb{R}^2$ projection, but we retain the depth coordinate ($q_i^B = f$), as this will be needed in formulating the epipolar constraint, below.

13. $\{f_j, f_k\}$ refer to the effective focal lengths corresponding to the $\{\hat{\jmath}, \hat{k}\}$ body dimensions. In a simple idealized pinhole camera, $f_j = f_k$, and we may dispense with the subscripts, as in Eq. 24.

14. Adiv G. Inherent ambiguities in recovering 3-D motion and structure from a noisy flow field. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989;11(5):477–489.

15. Taylor JR. Classical mechanics. University Science Books; 2005.

16. For simulation purposes, this condition is satisfied for all ϕ when the magnitude of the input perturbation angles $\{\theta_p, \psi_p\}_{max}$ are each less than $\cos^{-1}\left(\frac{1}{2^{1/4}}\right)$, or about 32.7°.

17. Video frames were rendered in *Mathematica*.

18. Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). Computer vision and image understanding. 2008:346–359.

19. In this simulation, the resulting constrained optimization problem was solved for using *Mathematica*'s global minimization function, NMinimize.

Nomenclature

$\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$	Body frame basis
$\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$	Earth frame basis
$\{x_0, y_0, z_0\}$	Initial projectile position, Earth frame (m)
$\{\theta_\mathbf{v}, \psi_\mathbf{v}\}$	Pitch and yaw angles associated projectile's translational velocity vector (rad)
$\{\theta_p, \psi_p\}$	Pitch and yaw perturbation angles (rad)
$\{\Gamma_E, \Gamma_B, \Gamma_{B'}\}$	Earth frame, inertial body frame, rotating body-frame
$\{\alpha, \beta\}$	Pitch and yaw angles of attack (rad)
$\{\phi, \theta, \psi\}$	Projectile euler angles: roll, pitch, and yaw; Earth frame (rad)
$\mathbf{b}_{\mathcal{O}}$	Body frame origin (projectile position) (m)
$\mathbf{e}_{\mathcal{O}}$	Earth frame origin (m)
$\mathbf{p}$	Scene point, $\mathbf{p} \in \mathbb{R}^3$ (m)
$\mathbf{q}$	Projected scene point, $\mathbf{q} \in \mathbb{R}^3$ (m)
$\mathbf{r}$	Projectile position vector (m)
$\mathbf{v}$	Projectile translational velocity vector (m/s)
$\boldsymbol{\omega}$	Projectile angular velocity (rad/s)
f	Focal length. $\{f_j, f_k\}$ refer to the effective focal lengths corresponding to the $\{\hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ body dimensions (m)
g	Acceleration of gravity (m/s^2)
t	Flight time (s)
Δt	Time laps between image frames (s)
FOV	Field of view (rad)
P	Projection matrix
QE	Launch angle above horizon (rad)
R	Rotation matrix
T	Translation matrix

INTENTIONALLY LEFT BLANK.